

Large language model (ChatGPT) as a support tool for breast tumor board

Year: 2023 | Citations: 166 | Authors: Vera Sorin, E. Klang, M. Sklair-Levy

Abstract

Large language models (LLM) such as ChatGPT have gained public and scientific attention. The aim of this study is to evaluate ChatGPT as a support tool for breast tumor board decisions making. We inserted into ChatGPT-3.5 clinical information of ten consecutive patients presented in a breast tumor board in our institution. We asked the chatbot to recommend management. The results generated by ChatGPT were compared to the final recommendations of the tumor board. They were also graded independently by two senior radiologists. Grading scores were between 1–5 (1 = completely disagree, 5 = completely agree), and in three different categories: summarization, recommendation, and explanation. The mean age was 49.4, 8/10 (80%) of patients had invasive ductal carcinoma, one patient (1/10, 10%) had a ductal carcinoma in-situ and one patient (1/10, 10%) had a phyllodes tumor with atypia. In seven out of ten cases (70%), ChatGPT's recommendations were similar to the tumor board's decisions. Mean scores while grading the chatbot's summarization, recommendation and explanation by the first reviewer were 3.7, 4.3, and 4.6 respectively. Mean values for the second reviewer were 4.3, 4.0, and 4.3, respectively. In this proof-of-concept study, we present initial results on the use of an LLM as a decision support tool in a breast tumor board. Given the significant advancements, it is warranted for clinicians to be familiar with the potential benefits and harms of the technology.